

学校编码: 10384
学 号: 23020061152457

分类号_____密级_____
UDC _____

厦 门 大 学

硕 士 学 位 论 文

用户定制口令声纹识别系统研究
与嵌入式实现

Research of User Customized Password Voiceprint
Recognition System and Embedded Implementation

陈燕

指导教师姓名: 洪青阳博士

专 业 名 称: 计算机应用

论文提交日期: 2009 年 月

论文答辩时间: 2009 年 月

学位授予日期: 2009 年 月

答辩委员会主席: _____

评 阅 人: _____

2009 年 月

厦门大学学位论文原创性声明

本人呈交的学位论文是本人在导师指导下,独立完成的研究成果。本人在论文写作中参考其他个人或集体已经发表的研究成果,均在文中以适当方式明确标明,并符合法律规范和《厦门大学研究生学术活动规范(试行)》。

另外,该学位论文为()课题(组)的研究成果,获得()课题(组)经费或实验室的资助,在()实验室完成。(请在以上括号内填写课题或课题组负责人或实验室名称,未有此项声明内容的,可以不作特别声明。)

声明人(签名):

年 月 日

厦门大学学位论文著作权使用声明

本人同意厦门大学根据《中华人民共和国学位条例暂行实施办法》等规定保留和使用此学位论文，并向主管部门或其指定机构送交学位论文（包括纸质版和电子版），允许学位论文进入厦门大学图书馆及其数据库被查阅、借阅。本人同意厦门大学将学位论文加入全国博士、硕士学位论文共建单位数据库进行检索，将学位论文的标题和摘要汇编出版，采用影印、缩印或者其它方式合理复制学位论文。

本学位论文属于：

（ ） 1. 经厦门大学保密委员会审查核定的保密学位论文，
于 年 月 日解密，解密后适用上述授权。

（ ） 2. 不保密，适用上述授权。

（请在以上相应括号内打“√”或填上相应内容。保密学位论文应是已经厦门大学保密委员会审定过的学位论文，未经厦门大学保密委员会审定的学位论文均为公开学位论文。此声明栏不填写的，默认为公开学位论文，均适用上述授权。）

声明人（签名）：

年 月 日

厦门大学博硕士论文摘要库

摘要

近年来,生物认证技术以其独特的方便性和经济性等优势受到世人瞩目,并日益成为人们日常生活和工作中的安全验证方式。声纹识别(又称说话人识别)是利用人体生物特征进行身份认证的一种技术,是目前最为方便与直接的识别技术。声纹识别根据说话内容分为文本相关和文本无关两种方式。

本文在研究声纹识别系统过程中,比较了文本无关方式和文本相关方式的性能,重点研究用户定制口令的声纹确认(UCP-SV)系统,并在嵌入式设备(PDA)上实现了语音录制、模型训练及声纹确认过程。

在研究声纹识别算法的过程中,本文主要有以下三个主要创新点:

- (1) 基于模型距离的 HMM 识别方法。首先对说话者进行模型训练,然后利用 HMM 模型距离来进行识别和确认。这种识别方法被成功地应用于司法鉴定中。
- (2) 采用动态阈值确认策略。声纹识别的识别过程中对于不同的人有不同的阈值,本文采取动态阈值的方法,对每个人分别训练出其阈值,相当于每个人有不同的确认参考。
- (3) 基于 HMM-UBM 二级判决识别方法。首先利用 UBM 这个代表大多数人的模型对说话人进行第一级的判决,将说话人分配到确定空间或不确定空间。若说话人被分配到不确定空间,再采用模型距离进行二级判决。该方法对于与说话人语音相似的冒认者有很好的拒绝效果。

在硬件平台上,本文使用的有PC平台,凌阳SPCE061A和PDA。在凌阳SPCE061A上成功地实现了嵌入式语音识别系统,并投入到实际应用中。本文还实现了基于模式匹配方法的声纹确认系统。但由于凌阳SPCE061A 硬件的存储和计算能力有限,效果欠佳,我们最终选择了有代表性的PDA(基于ARM芯片)作为嵌入式实现平台。在实现过程中重点阐述了在PDA上的计算优化:浮点转定点,计算预处理,快速近似计算等方法,从而提高系统运算效率,降低运算负荷,以保证嵌入式实时性的要求。

关键词: 声纹识别; 用户定制口令; HMM-UBM。

Abstract

The technology of voiceprint recognition (speaker recognition) has the advantages of convenience, benefit and accuracy. It is used more and more popular in daily life. Voiceprint recognition is a kind of authentication based on biological features. According to the text content, it can be divided to the types of text-dependent and text independent.

In this work, we focus on User Customized Password (UCP) voiceprint recognition system. We have implemented it on PDA, which includes voice collection, model training, and voiceprint verification.

In this work, we have proposed three novel methods:

1. HMM recognition methods based on model distance. Model distance stands for the small difference between speakers. First, we train the speaker models in advance, and then carry out the verification process using model distance. This method has been adopted successfully in Forensic Authentication.

2. Dynamic threshold for the verification. During the verification process, different people may have different thresholds. In this work, we adopt dynamic threshold and train different thresholds for different speakers.

3. Two-level decision mechanism based on HMM-UBM. In this method, the first decision is made according to the UBM. After the first step, the speaker is located in the certain or uncertain space. If the speaker is in the uncertain space, then the second decision will be conducted based on model distance according to dynamic threshold.

For the hardware platform, personal computer, Sunplus SPCE061A and PDA are all used. We have successfully developed embedded speech recognition system in SPCE061A and have applied it in practice. However, when developing speaker recognition on this platform, the system doesn't perform ideally for its limitation of storage and calculation. To have a better performance, we design and implement the speaker verification system based on PDA. Several new methods are applied to

improve the verification performance, such as the transformation from floating-point operation to fixed-point operation, pretreatment of the calculation, fast and approximate mathematic calculation etc. These methods guarantee the real-time performance of embedded system.

Key Words: Voiceprint Recognition; User Customized Password; HMM-UBM.

目 录

第一章 绪论	1
1.1 声纹识别概述	1
1.2 声纹识别的历史	3
1.3 嵌入式声纹识别的现状	4
1.4 嵌入式设备上声纹识别的难点	5
1.5 PC、凌阳与 PDA 实验平台	6
1.6 论文的主要工作及组织结构	8
第二章 声纹识别基础及性能评价标准	9
2.1 语音基本特征	9
2.1.1 LPCC 特征参数的提取	10
2.1.2 MFCC 特征参数的提取	12
2.2 声纹识别模型概述	14
2.2.1 模型匹配法	15
2.2.2 概率统计方法	15
2.2.3 辨别分类器方法	16
2.3 文本无关的 GMM-UBM 声纹识别	17
2.4 声纹识别的性能评价标准	19
2.4.1 声纹辨认评价标准	19
2.4.2 声纹确认评价标准	19
2.5 小结	22
第三章 用户定制口令 HMM 模型距离方法	23
3.1 基于 HMM 模型距离方法	23
3.2 动态阈值声纹识别策略	25
3.3 HMM 模型距离确认方法在司法鉴定上的应用	27
3.3.1 司法鉴定声纹识别基本原理	27
3.3.2 声纹识别系统验证	28
3.3.3 检验材料测试结果	30

3.3.4 鉴定结果分析.....	31
3.4 小结.....	31
第四章 用户定制口令 HMM-UBM 二级判决方法.....	33
4.1 HMM-UBM 二级判决方法	33
4.1.1 HMM-UBM	33
4.1.2 二级判决机制.....	34
4.2 声纹识别方法实验结果与比较	35
4.2.1 GMM-UBM 与用户定制口令 HMM-UBM 二级判决比较.....	35
4.2.2 HMM 模型距离与 HMM-UBM 二级判决比较	36
4.2.3 UBM 混合数对 HMM-UBM 方法的影响	37
4.3 小结.....	38
第五章 凌阳 SPCE061A 板声纹识别系统的实现.....	39
5.1 SPCE061A 单片机简介	39
5.1.1 61 板接口说明.....	39
5.1.2 精简 61 板及其在嵌入式语音识别上的应用.....	41
5.2 SPCE061A 板声纹识别系统的实现	43
5.2.1 系统设计.....	43
5.2.2 系统测试.....	45
5.3 小结.....	46
第六章 PDA 上的声纹识别系统的实现.....	47
6.1 PDA 简介	47
6.2 PDA 上声纹识别系统的计算优化	48
6.2.1 数据格式.....	48
6.2.2 四则运算的转换.....	50
6.2.3 计算的预处理.....	50
6.2.4 数学的快速近似计算.....	51
6.3 PDA 上系统的实现	51
6.3.1 特征提取.....	52

6.3.2 模型的训练及更新.....	53
6.3.3 结果判断.....	54
6.3.4 实验结果.....	54
6.4 小结.....	56
第七章 总结与展望	57
参考文献.....	59
发表论文.....	63
致 谢.....	65

Contents

Chapter 1 Introduction.....	1
1.1 Brief introduction of Voiceprint Recognition	1
1.2 The history of Voiceprint Recognition	3
1.3 The present research of Embedded Voiceprint Recognition.....	4
1.4 The difficulties of Embedded Voiceprint Recognition	5
1.5 Experiment platforms.....	6
1.6 Research contents and chapter organization.....	8
Chapter 2 Voiceprint Recognition and Evaluation Standards	9
2.1 Basic speech features	9
2.1.1 The extraction of LPCC	10
2.1.2 The extraction of MFCC	12
2.2 The introduction of models for Voiceprint Recognition	14
2.2.1 Pattern Matching	15
2.2.2 Stochastic	15
2.2.3 Classification	16
2.3 Text-independent GMM-UBM	17
2.4 Evaluation standards for Voiceprint Recognition.....	19
2.4.1 Evaluation standards for Speaker Identification	19
2.4.2 Evaluation standards for Speaker Verification	19
2.5 Conclusion	22
Chapter 3: UCP Model distance based on HMM	23
3.1 Model distance based n HMM	23
3.2 Dynamic threshold for recognition.....	25
3.3 Model distance based on HMM-UBM used in Forensic Authentication	27
3.3.1 Theory for Forensic Authentication	27
3.3.2 Testing of Speaker Verification system.....	28

3.3.3 Result of juridic material	30
3.3.4 Result analysis.....	31
3.4 Conclusion	31
Chapter 4 UCP Two-level decision rule based on HMM-UBM.....	33
4.1 Two-level decision rule based on HMM-UBM	33
4.1.1 HMM-UBM.....	33
4.1.2 Two-level decision rule	34
4.2 The comparison of methods for Voiceprint Recognition.....	35
4.2.1 The comparison between GMM-UBM and UCP HMM-UBM.....	35
4.2.2 The comparison between model distance based on HMM and UCP HMM-UBM.....	36
4.2.3 The influence of mixture numbers for UCP HMM-UBM	37
4.3 Conclusion	38
Chapter 5: Voiceprint Recognition system on SPCE061A.....	39
5.1 The introduction of Sunplus SPCE061A	39
5.1.1 The interface of Sunplus SPCE061A	39
5.1.2 The application of Simplified Sunplus SPCE061A in Speech Recognition.....	41
5.2 The implementation of Voiceprint Recognition on SPCE061A	43
5.2.1 System design	43
5.2.2 System evaluation.....	45
5.3 Conclusion	46
Chapter 6: Voiceprint Recognition system on PDA	47
6.1 The introduction of PDA	48
6.2 Arithmetic optimization for the Voiceprint Recognition system	48
6.2.1 Data format.....	48
6.2.2 The transformation of operation	50
6.2.3 Arithmetic pretreatment	50
6.2.4 Fast and approximate operation.....	51

6.3 The implementation of system on PDA.....	51
6.3.1 Feature extraction	52
6.3.2 Model training and updating	53
6.3.3 Decision.....	54
6.3.4 Result	54
6.4 Conclusion	56
Chapter 7: Conclusion and Future work.....	57
References	59
Published Papers.....	63
Acknowledgement	65

厦门大学博硕士论文摘要库

第一章 绪论

当今信息社会中，在国家安全、金融、执法司法，公司等社会各个领域均需要个人身份验证。近年来，在生物认证技术领域，声纹识别(又称说话人识别)技术以其独特的方便性、经济性和准确性等优势受到世人瞩目，并日益成为人们日常生活和工作中重要且普及的安全验证方式。和指纹识别、虹膜识别一样，声纹识别属于生物识别的一种，它是一项根据语音波形中反映说话人生理和行为特征的语音参数，自动识别说话人身份的技术。说话人识别是交叉运用心理学、生理学、语音信号处理、模式识别、统计学习理论和人工智能的综合性研究课题^[1]。

适合生物识别的特征应符合以下几个性质：通用性(每个人都有)、独特性(每个人都不一样)、恒定性(不随时间而变化)、可收集性(可定量测量)、准确性(随量的增加而降低)、高可采用性(人们易接受)和低欺骗性(不易被欺骗)。人的语音不仅具备以上性质，跟其他的生物测定技术相比还有很大的优势。人的语音可以很自然的产生，这使得说话人识别具有其他生物认证技术不具备的优势：(1)数据易取性；(2)采集成本低廉性；(3)远程的可操作性；(4)嵌入式系统的可行性；(5)身份认证的隐蔽性。这些优势使得说话人识别的应用越来越受到系统开发者和用户的青睐，可以说说话人识别几乎可以应用到我们日常生活的各个角落，如信息领域，银行、证券、公安司法、国防、国家安全及证件防伪等。据统计，说话人识别的世界市场占有率为 15.8%，仅次于手指和手的生物特征识别，并有不断上升的趋势。

1.1 声纹识别概述

声纹识别(Voiceprint Recognition)是由说话人识别(Speaker Recognition)引出的，两者其实是同一个概念，本文统一用声纹识别来表述。最早由 Bell 实验室的 L.G. Kersta 目视观察语谱图进行识别，提出了“声纹(Voiceprint)”的概念。声纹识别按其最终完成的任务可分为两类：声纹辨认(Speaker Identification, SI)和声纹识别(Speaker Verification, SV)，如图 1.1 所示，前者是根据说话人语音确

定为 N 个参考说话人中的某一个，是一个选择问题，声纹辨认系统的性能将随着说话人集合的规模增大而降低；后者是证实说话人的身份与其声明的是否一致，是一个二选一的判定问题，声纹识别系统的性能与说话人集合的规模无关。

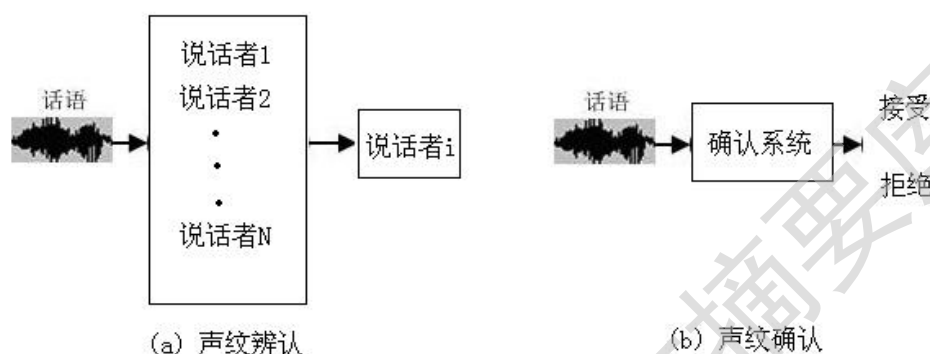


图1.1：声纹识别分类

根据测试时所采用文本是否与训练时所采用文本相同，声纹识别系统可分为文本相关(Text-Dependent)和文本无关(Text-Independent)两种方式。与文本有关的说话人识别系统要求用户按照规定的内容发音，每个人的声纹模型逐个被精确地建立，而识别时也必须按规定的内容发音，因此可以达到较好的识别效果。而与文本无关的识别系统则不规定说话人的发音内容，模型建立相对困难，但用户使用方便，可应用范围较宽。图 1.2 是声纹识别系统的典型结构^[2]。

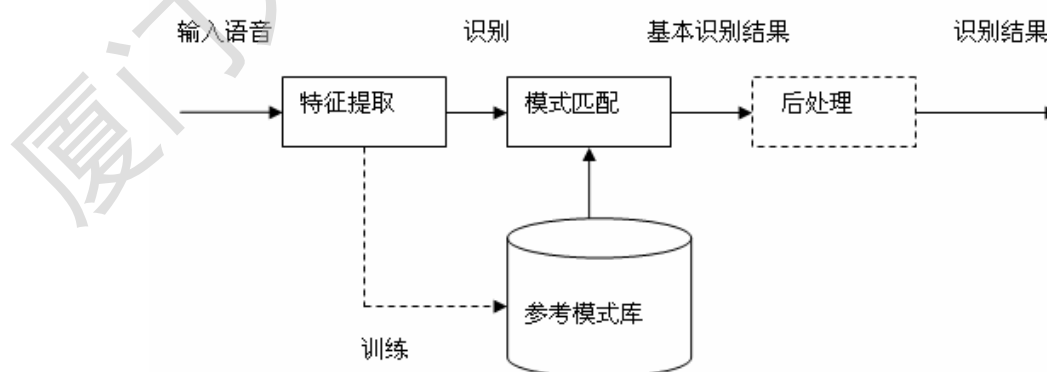


图1.2：声纹识别系统的典型结构

该过程包括训练和识别两个阶段，训练时每个人重复一定次数的语音，提

Degree papers are in the "[Xiamen University Electronic Theses and Dissertations Database](#)". Full texts are available in the following ways:

1. If your library is a CALIS member libraries, please log on <http://etd.calis.edu.cn/> and submit requests online, or consult the interlibrary loan department in your library.
2. For users of non-CALIS member libraries, please mail to etd@xmu.edu.cn for delivery details.

厦门大学博硕士论文摘要库